(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number : **0 685 800 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number : **95303586.2**

(22) Date of filing : **26.05.95**

(51) Int. Cl.[6] : **G06F 17/10**

(30) Priority : **28.05.94 JP 138054/94**
28.05.94 JP 138055/94

(43) Date of publication of application :
06.12.95 Bulletin 95/49

(84) Designated Contracting States :
DE FR GB

(71) Applicant : **SONY CORPORATION**
7-35 Kitashinagawa 6-chome
Shinagawa-ku
Tokyo 141 (JP)

(72) Inventor : **Kondo, Tetsujiro, c/o Intellectual Property Div.**
Sony Corp.,
6-7-35 Kitashinagawa
Shinagawa-ku, Tokyo 141 (JP)

(74) Representative : **Williams, Janice et al**
D. Young & Co.,
21 New Fetter Lane
London EC4A 1DA (GB)

(54) Method and apparatus for digital signal processing.

(57) In processing a digital information signal (a digital video signal etc), an improvement in the accuracy of the signal process can be made by integrating the time region process and the frequency region process. In one embodiment, an input digital video signal is supplied to a DCT circuit 3. The DCT circuit 3 converts the input video digital signal into coefficient data. A categorizing circuit 5 separates a flat component 6a and an impulse component 6b in a frequency region from the input signal. The flat component 6a is supplied to an inverse DCT circuit 7. The inverse DCT circuit 7 converts the input signal into a signal on a time axis. A class categorizing adaptive processing circuit 9 compensates a resolution in a time region. The impulse component 6a is supplied to a gain converting circuit 10. The gain converting circuit 10 compensates the high band in a frequency region. The output signal of the gain converting circuit 10 is supplied to an inverse DCT circuit 11. The inverse DCT circuit 11 converts the input signal into a signal on a time axis. A mix circuit 14 mixes signals that have been processed. Thus. a video signal compensated the resolution is obtained from an output terminal 15. The adaptive procession circuit 9 and the gain converting circuit 10 use pre-learnt predicted coefficients, predicted values, and gain converting ratio.

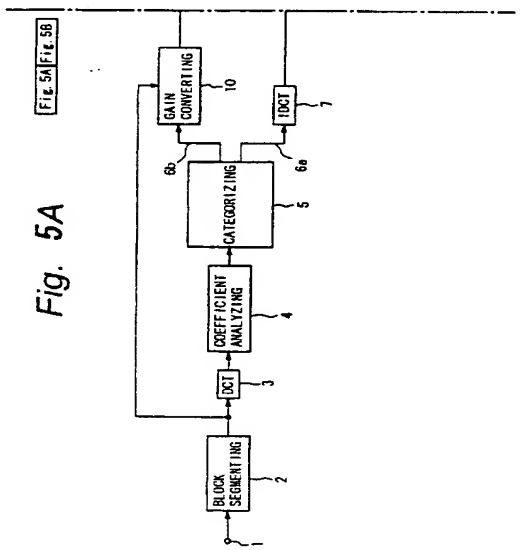
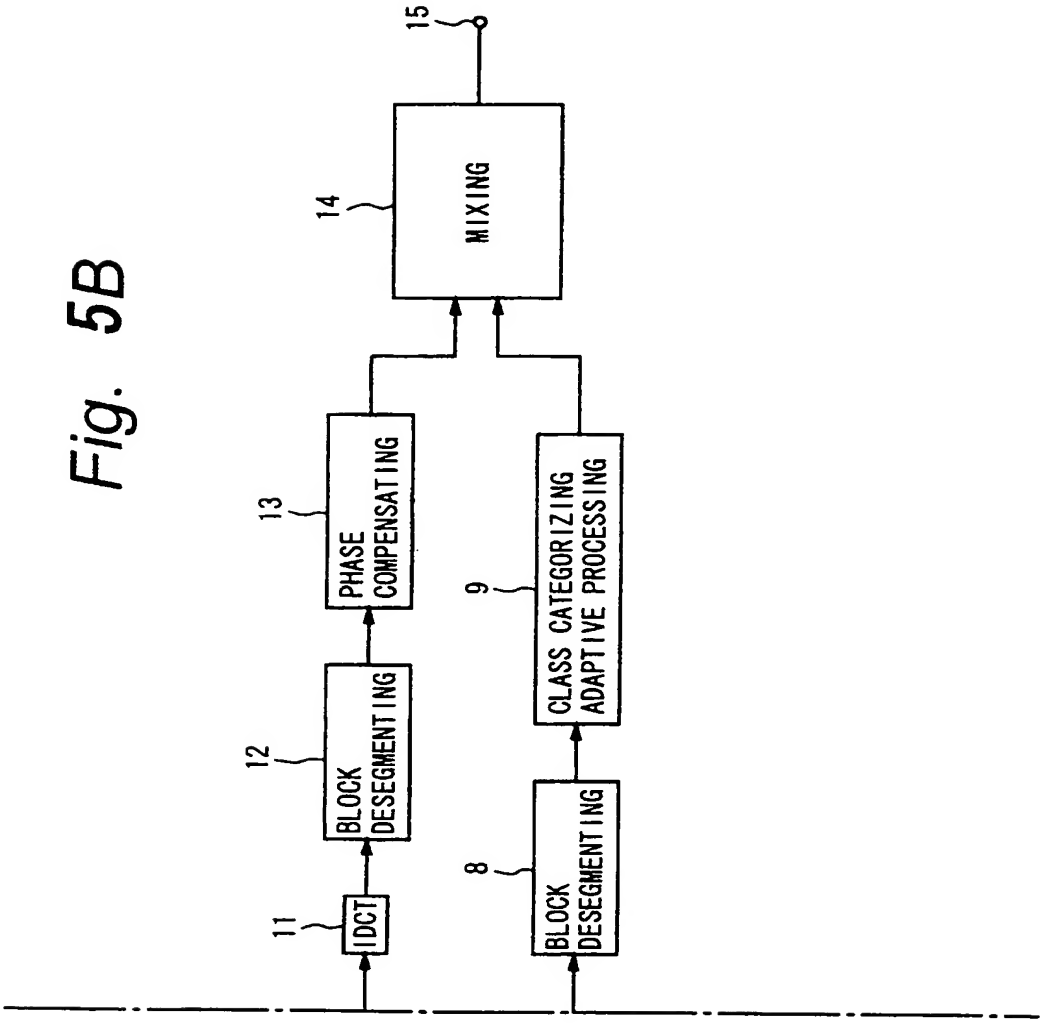


Fig. 5A

EP 0 685 800 A2

Fig. 5B


15
14
MIXING
13
PHASE COMPENSATING
9
CLASS CATEGORIZING ADAPTIVE PROCESSING
12
BLOCK DESEGMENTING
8
BLOCK DESEGMENTING
11
IDCT

The present invention relates to signal processing apparatus and methods for digital signals such as a digital image signal and a digital audio signal.

Conventional signals such as digital image signal and digital audio signal are processed in a time region or in a frequency region. In the process of the frequency region, the component with the steady characteristic (the steady component) can be properly represented, while the component with the transient characteristic (the transient component) cannot be properly represented. In contrast, in the process of the time region, the component with the transient characteristic can be properly represented, while the component with the steady characteristic cannot be properly represented. The steady characteristic represents a stable repetitive change. On the other hand, the transient characteristic represents an independent single change.

As an example, Fig. 1A and 1B show the case of a process in the time region. As shown in Fig. 1A, the component having the transient characteristic represents a wave form that largely changes on the time axis (impulse wave form). After sampling the wave form, it can be satisfactorily processed with for example several samples. Dots in the wave form represent sampling points. In the case of a digital signal, the dots represent a discrete signal sequence with sampling values corresponding to the levels of sampling points. However, in the case shown in Fig. 1A and 1B, the transient characteristic can be represented by an analog signal wave. This applies to the following description. On the other hand, the component with the steady characteristic represents a wave form that gradually varies (flat wave form) on the time axis as shown in Fig. 1B. In this case, the feature of the wave form cannot be obtained with several samples. Thus, the steady characteristic cannot be satisfactorily processed.

In the frequency region, since the component with the steady characteristic includes one kind of single component or a few kinds of frequency components, it is represented as an impulse wave form as shown in Fig. 2A. On the other hand, the component with the transient characteristic is represented as a flat wave form as shown in Fig. 2B. As with the case of the time region, the impulse wave form more suitably represents the feature of the signal than the flat wave form.

A general signal wave form includes the component with the steady characteristic (flat) portions FL1, FL2, FL3, ... etc. and the component with the transient characteristic (impulse) portions IM1, IM2, ... etc. on the time axis as shown in Fig. 3. Thus, when one of the time region process and the frequency region process is performed, the feature of the signal cannot be correctly obtained. Thus, for the same signal, the process in the time region and the process in the frequency region should be performed, thereby increasing the process time and/or hardware scale.

According to an aspect of the invention, there is provided a digital signal processing apparatus, comprising at least one analyzing means of first analyzing means for analyzing an input digital signal in a time region and second analyzing means for analyzing the input digital signal in a frequency region, a categorizing means for categorizing the input digital signal into two components corresponding to an analyzed output of the analyzing means, first and second processing means for adaptively processing the signals categorized by the categorizing means in the time region and in the frequency region, respectively, and a mixing means for mixing outputs of the first and second processing means.

Respective further aspects of the invention are set forth in claims 2, 10. 11, 12, 13, 29 and 30.

Embodiments of the present invention provide a digital signal processing apparatus for processing the component with the steady characteristic portions of a digital signal in a frequency region and the component with the transient characteristic portions of the digital signal in a time region while decreasing the process time and hardware scale.

In embodiments of the invention, an input digital signal (for example. a digital video signal) is transformed by DCT into coefficient data. The coefficient data is analyzed so as to separate an impulse component and a flat component in a frequency region from the coefficient data. In the frequency region, the impulse signal is processed by a first processing means. The flat component in the frequency region becomes an impulse signal in a time region. Thus. the impulse component in the frequency region is processed by a second processing means in the time region. Since the signal is processed in the impulse shape. the result of the process is improved.

Embodiments of the invention will now be described. by way of example. with reference to the accompanying drawings in which:

Fig. 1A and 1B are schematic diagram showing an impulse component and a flat component in a time domain;

Fig. 2A and 2B are schematic diagram showing an impulse component and a flat component in a frequency region;

Fig. 3 is a schematic diagram showing a signal wave form including both an impulse component and a flat component in a time region;

Figs. 4A and 4B are schematic diagram for explaining the compensation of resolution according to an em-

bodiment of the present invention;

Fig. 5 (5A and 5B) is a overall block diagram of an embodiment of the present invention;

Fig. 6 is a block diagram showing an example of a class categorizing adaptive processing circuit according to an embodiment of the present invention;

Fig. 7 is a schematic diagram showing an arrangement of pixels between SD pixels and HD pixels;

Fig. 8 is a block diagram showing an example of the construction for generating a mapping table that stores predicted coefficients;

Fig. 9 is a block diagrams showing an example of the construction for generating a mapping table that stores predicted values;

Fig. 10 is a block diagram showing an example of the construction for forming predicted coefficients or predicted values in learning state;

Fig. 11 is a schematic diagram showing another example of an arrangement of pixels between SD pixels and HD pixels;

Fig. 12 is a flow chart showing a process for forming predicted coefficients in learning state;

Fig. 13 is a flow chart showing a process for forming predicted values in learning state;

Fig. 14 is a schematic diagram for explaining a high band compensation in a frequency region; and

Fig. 15 is a block diagram for learning a gain converting ratio for high band compensation in a frequency region.

The embodiment to be described is applied for the compensation of resolution of a digital video signal. As shown in Fig. 4A, when the resolution is compensated, a band (with frequency characteristic 20a) of a video signal that is narrowed by a filtering process or the like is compensated into a band (with frequency characteristic 20b). In other words, by generating a component represented by a hatched portion, a narrow band of the video signal is converted into a wide band of the video signal as shown in Fig. 4B.

Fig. 5 shows the overall construction of the embodiment. In Fig. 5, a digital video signal with a standard resolution (referred to as an SD video signal) is supplied to an input terminal 1. A digital video signal with a high resolution is referred to as an HD video signal. Examples of the input SD video signal are an SDVCR (Standard Definition Video Cassette Recorder) reproduction signal, broadcasting signal, and so forth. The input SD video signal is supplied to a block segmenting circuit 2. The block segmenting circuit 2 scans and converts a video signal in TV raster sequence into a signal with a block structure of (8 x 8).

A DCT (Discrete Cosine Transform) circuit 3 is connected to the block segmenting circuit 2. The DCT circuit 3 generates coefficient data DC of one DC component and coefficient data AC1, AC2, ..., AC63 of 63 AC components. As an example, the coefficient data are output in zigzag scanning method from DC coefficient data to higher order AC coefficient data. The DCT is one means for analyzing the frequency of the input video signal. Alternatively, FFT (Fast Fourier Transformation), Hadamard transformation, and so forth may be used.

The coefficient data that are output from the DCT circuit 3 are supplied to a categorizing circuit 5 through a coefficient analyzing circuit 4. The coefficient analyzing circuit 4 and the categorizing circuit 5 separate a steady component and a transient component from the digital video signal that has been converted into the frequency region. The categorizing circuit 5 outputs a flat component in the frequency region (namely, the transient component) 6a and an impulse component (namely, the steady component) 6b.

For easy understanding, it is assumed that values of coefficient data are DC = 50, AC1 = 48, AC2 = 46, AC3 = 44, AC4 = 42, AC5 = 60, and so forth. The coefficient analyzing circuit 4 analyzes the coefficient data and determines that AC5 is an impulse component. In other words, due to the tendency of changes of AC1, AC2, AC3, and AC4, AC5 must be 40. However, since ACS is actually 60, it deviates for 20. The categorizing circuit 5 separates flat components in the frequency region (transient component, in the example, DC = 50, AC1 = 48, AC2 = 46, AC3 = 44, AC4 = 42, AC5 = 40, ...) 6a and pulse components in the frequency regions (steady components, in the example, DC = 0, AC1 = 0, AC2 = 0, AC3 = 0, AC4 = 0, AC5 = 20, ...) 6b.

The flat component 6a that is output from the categorizing circuit 5 is supplied to an inverse DCT circuit 7. The inverse DCT circuit 7 restores a signal in the time region. The signal in the time region is supplied to a block desegmenting circuit 8. The block desegmenting circuit 8 outputs a digital video signal in the TV raster scanning sequence. The digital video signal is supplied to a class categorizing adaptive processing circuit 9 as a second processing circuit. As will be described later, the circuit 9 increases the resolution in the time region. The flat component 6a is suitable for the process in the time region. Thus, the circuit 9 properly compensates the resolution.

The impulse component 6b that is output from the categorizing circuit 5 is supplied to a gain converting circuit 10. The output signal of the block segmenting circuit 2 is supplied to the gain converting circuit 10 so as to categorize the output signal as a class. The gain converting circuit 10 has a memory that stores pre-learnt gain converting ratio information as will be described later. By adjusting the gain of the coefficient data corresponding to the converting ratio information, the high band component in the frequency region is increased.

The output signal of the gain converting circuit 10 is supplied to an inverse DCT circuit 11. The inverse DCT circuit 11 restores a signal in the time region. The signal in the time region is supplied to a block desegmenting circuit 12. The block desegmenting circuit 12 converts the signal in the time region into data in the TV raster scanning sequence.

An output signal of the block desegmenting circuit 12 is supplied to a mixing circuit 14 through a phase compensating circuit 13. The mixing circuit 14 mixes the output signal of the block desegmenting circuit 12 and the output signal of the class categorizing adaptive processing circuit 9. The mixing circuit 14 simply mixes these signals. The mixing circuit 14 outputs a digital video signal whose resolution has been compensated (namely, an HD video signal) to an output terminal 15.

Fig. 6 shows an example of the class categorizing adaptive processing circuit 9. A digital video signal that is output from the block desegmenting circuit 8 is supplied to an input terminal 21. The digital video signal is a flat component (transient component) of the SD video signal that is an impulse component in the time region. The digital video signal is supplied to a simultaneous processing circuit 2. The output data of the simultaneous processing circuit 22 is supplied to a class categorizing circuit 23. The output of the class categorizing circuit 23 is supplied to memories 24a to 24d as an address signal. The memories 24a to 24d store mapping tables M1 to M4, respectively.

Fig. 7 shows a partial schematic diagram showing the relation-between an SD image and an HD image. In Fig. 7, pixel data denoted by O represent the SD image. Pixel data denoted by X represent the HD image. For example, four HD pixel data y1 to y4 are composed of 12 SD pixel data a to 1 by the class categorizing adaptive processing circuit 9. The mapping table M1 of the memory 24a generates the pixel data y1. The mapping tables M2, M3, and M4 of the memories 24b, 24c, and 24d generate the pixel data y2, y3, and y4, respectively.

The outputs that are read from the memories 24a to 24d are supplied to a selector 25. The selector 25 is controlled by the output of a select signal generating circuit 26. A sample clock of an HD image is supplied to the select signal generating circuit 26 through an input terminal 27. The selector 25 selects the four pixel data y1 to y4 one after the other. These pixel data are supplied to a scanning converting circuit 28. The scanning converting circuit 28 generates and outputs the pixel data of the HD image to an output terminal 29 in the raster scanning sequence. The number of pixels of an output image is four times larger than the number of pixels of an input SD video signal.

The mapping tables M1 to M4 stored in the memories 24a to 24d are pre-learnt. Fig. 8 shows an example of the construction for generating the mapping tables M1 to M4. In Fig. 8, a digital HD video signal is supplied to an input terminal 31. The HD video signal is preferably a standard signal considered for generating the mapping tables. In reality, a standard image is photographed by an HD video camera. By recording the photograph signal to an HDVTR, an HD video signal is obtained.

The HD video signal is supplied to a simultaneous processing circuit 32. The simultaneous processing circuit 32 simultaneously outputs pixel data a to 1 and y_2 to y_4 that represent the relation of positions shown in Fig. 7. The pixel data a to 1 are supplied to a class categorizing circuit 33. The class categorizing circuit 33 categorizes the HD pixel data y_1 to y_4 as classes corresponding to tone, pattern, or the like. The output of the class categorizing circuit 33 is supplied to mapping table generating circuits 34a to 34d in common.

The pixel data y_1 to y_4 that are output from the simultaneous processing circuit 32 are supplied to the mapping table generating circuits 34a to 34d, respectively. The structure of each of the mapping table 34a to 34d is the same. There are two types of mapping tables. One type is used to predict the values y_1, y_2, y_3, and y_4 of the HD pixels by linear combination of the values a to 1 of the SD pixels and the coefficients w_1 to w_{12}. In this case, the coefficients w_1 to w_{12} are obtained for each class. The other type is used to obtain the values of the HD pixels predicated for each class.

The memories with which the mapping table generating circuits 34a to 34d shown in Fig. 8 are provided store mapping tables that represent the correlation between the HD video signal and the SD video signal. In other words, when a plurality of data of the SD video signal are provided, a mapping table that outputs the pixel data of the HD video signal that averagely corresponds to the class of the data can be formed.

As with the class categorizing circuit 23 shown in Fig. 6, the class categorizing circuit 33 categorizes considered pixel data as a class and generates class information. The considered pixel can be categorized as a class corresponding to tone, pattern, or the like. When the tone is used and the number of pixel data is composed of eight bits, since the number of classes becomes very large, the number of bits of each pixel is preferably decreased by highly efficient encoding method such as ADRC method. A plurality of patterns each of which is composed of four pixels (for example, flat pattern and patterns that increase in upper right direction and decrease in lower left direction, etc.) are provided. The output data of the simultaneous processing circuit 32 is categorized as one of a plurality of patterns.

For example, the mapping table generating circuit 34a that obtains the HD pixel data y_1 has a memory to

which class information that is an address is supplied from the class categorizing circuit 33. When the training (learning) operation is performed, the original HD video signal is thinned out so as to form an SD video signal. In this case, horizontal decimation process (for sub samples) and the vertical decimation process (for sub lines) are performed. An HD video signal for one frame or more (for example, a still image) is used. Sample values of pixel data a to 1 and yl are written to addresses of the memory corresponding to class information. For example, $(a_{10}, a_{20}, ..., a_{n0})$ $(b_{10}, b_{20}, ..., b_{n0})$ ... $(1_{10}, 1_{20}, ..., 1_{n0})$ (y_{10}, y_{20}, y_{n0}) are stored in the address ADO of the memory.

The learnt data is read from the memory. A coefficient that minimizes the error between the predicted value of the HD pixel (corresponding to yl) obtained by liner combination of the values a to 1 of the SD pixels and coefficients w_1 to w_{12} and the true value is obtained by the method of least squares. When learnt data stored at the address of the memory is considered, the following simultaneous equations with respect to learnt data stored at an address of one memory are satisfied.

$$y_{10} = w_1a_{10} + w_2b_{10} + w_3c_{10} + ... + w_{12}l_{10}$$
$$y_{20} = w_1a_{20} + w_2b_{20} + w_3c_{20} + ... + w_{12}l_{20}$$
$$y_{30} = w_1a_{30} + w_2b_{30} + w_3c_{30} + ... + w_{12}l_{30}$$
$$y_{n0} = w_1a_{n0} + w_2b_{n0} + w_3c_{n0} + ... + w_{12}l_{n0}$$

Since y_{10} to y_{n0}, a_{10} to a_{n0}, b_{10} to b_{n0}, c_{10} to c_{n0}, ..., 1_{10} to 1_{n0} are known, coefficients w_1 to w_{12} that minimize the squares of errors of predicted values for y_{10} to y_{n0} (true values) can be obtained. The coefficients of other classes (addresses) can be obtained in the same manner. The coefficients obtained in this manner are stored in the memory and used as a mapping table.

In addition to the coefficients, the values of data of the HD video signal for each class may be obtained by training and then stored in the memory. Fig. 9 shows the construction for performing such an operation. In Fig. 9, a data memory 40 and a counting memory 41 are provided. The class information is supplied as an address to the data memory 40 and the counting memory 41.

The output that is read from the couting memory 41 is supplied to an adding device 42. The adding device 42 increments the value that is output from the counting memory 41 by +1. The output of the adding device 42 is written to the same address of the memory 41. As an initial state, the content of each address of the memories 40 and 41 is cleared to zero.

The data that is read from the data memory 40 is supplied to a multiplying device 43 and multiplied by the count value read from the counting memory 41. The output of the multiplying device 43 is supplied to an adding device 45. The adding device 45 adds the input from the multiplying device 43 and the input data y. The output of the adding device 44 is supplied as a divisor to the dividing device 45. The output (quotient) of the dividing device 45 is input as input data to the data memory 40.

In the construction shown in Fig. 9, when a particular address is accessed at the first time, since the outputs of the memories 40 and 41 are 0, the data y_{10} is directly written to the memory 40. The value of the address corresponding to the memory 41 becomes 1. Thereafter, when the address is accessed again, the output of the adding device 42 is 2 and the output of the adding device 44 is $(y_{10} + y_{20})$. Thus, the output of the dividing device 45 is $(y_{10} + y_{20}) / 2$. This output is written to the memory 40. When the above-described address is accessed, in the same operation, the data of the memory 40 is changed to $(y_{10} + y_{20} + y_{30}) / 3$ and the count number is updated to 3.

When the above-described operation is performed for a predetermined period, a class is designated to the memory 40 corresponding to the output of the class categorizing circuit 33. A mapping table that outputs such data is stored in the memory 40. In other words, the mapping table that outputs data averagely corresponds to the class categorized against a plurality of pixel data of the input video signal is formed.

Next, the class categorizing adaptive processing circuit 9 will be described in detail. The class categorizing adaptive processing circuit 9 predetermines the coefficients of the linear combination by training operation. When the training operation is performed, the construction shown in Fig. 10 is used. In Fig. 10, reference numeral 51 is an input terminal. A plurality of still images of standard HD signals are input to the input terminal 51. A vertical decimation filter 52 thins out each input HD image by 1/2. The vertical decimation filter 52 is connected to a horizontal decimation filter 53. The horizontal decimation filter 53 thins out each input HD image by 1/2. Still images with the same number of pixels as the SD signal are supplied to a learning portion 54. A memory 55 stores the class code generated by the learning portion 54 and the learnt result.

In this example, as shown in Fig. 11, the relation of positions of HD pixels and SD pixels is defined. As shown in Fig. 11, when an SD pixel (3 x 3) block is used, one set of SD pixels a to i and HD pixels A, B, C, and D becomes learnt data. When there are a plurality of sets of learnt data for each frame and the number of frames is increased, a large number of sets of learnt data can be used.

Fig. 12 is a flow chart showing the operation of software for the learning portion 54 that determines coefficients of linear combination. At step 61, the control of the learning portion is started. At step 62 "data block

segmentation", an HD signal and an SD signal are supplied and a process for extracting HD pixels and SD pixels that have the relation shown in Fig. 11 is performed. At step 63 "data end", when the process for all data (for example, one frame) has been finished, the flow advances to step 66 "determining predicted coefficients". When the process for one frame has not been finished, the flow advances to step 64 "determining class".

At step 64 "determining class", a class is determined corresponding to the signal pattern of the SD signal. In this control, ADRC can be used so as to reduce the number of bits. At step 65 "adding normal equations", equations that will be described later are generated.

After all data have been processed at step 63 "data end", the flow advances to step 66. At step 66 "determining predicted coefficients", equations that will be described later are solved corresponding to matrix solution so as to determine predicted coefficients. At step 67 "storing predicted coefficients", the predicted coefficients are stored in the memory. At step 68, the control of the learning portion is finished. In the memory, the predicted coefficients of the class that is determined as an address corresponding to the SD signal are stored. The class and the predicted coefficients correspond to the above-described mapping table.

Next, the process for obtaining coefficients that define the relation between HD pixels and SD pixels shown in Fig. 11 will be described in detail. Now, it is assumed that the SD pixel levels are denoted by x_1 to x_n and the HD pixel level is denoted by y. In this case, a linear estimating equation with n taps of coefficients w_1 to w_n for each class is represented as follows.

$$y' = w_1x_1 + w_2x_2 + \ldots + w_nx_n \quad (1)$$

Before the learning operation is performed, w_i is an unknown coefficient.

As described above, a plurality of HD data and SD data are learnt for each class. When the number of data is m, the following expression is obtained corresponding to the equation 1.

$$y_j' = w_1x_{j1} + w_2x_{22} + \ldots + w_nx_{jn} \quad (2)$$

(where j = 1, 2, ... m)

When m > n, since w_1 to w_n are not uniquely determined, elements of the error vector e are defined as follows and coefficients that minimize the value of the equation 4 are obtained.

$$e_j = y_j - (w_1x_{j1} + w_2x_{j2} + \ldots + w_nx_{jn}) \quad (3)$$

(where j = 1, 2, ... , m)

$$E^2 = \sum_{j=0}^{m} \{e_j\}^2 \quad (4)$$

In other words, the coefficients are obtained by so-called method of least squares. Next, partial differential coefficients are obtained with respect to w_i of the equation 3.

$$\frac{\partial E^2}{\partial w_i} = \sum_{j=0}^{m} 2\left(\frac{\partial e_j}{\partial w_i}\right)e_j = \sum_{j=0}^{m} 2x_{ji} \cdot e_j \quad (5)$$

Since each w_i is obtained so that the value of the equation 6 becomes 0,

$$X_{ji} = \sum_{p=0}^{m} x_{pi} \cdot x_{pj} \quad (6)$$

$$Y_i = \sum_{j=0}^{m} x_{ji} \cdot y_j \quad (7)$$

The equations 6 and 7 can be represented by the following matrix.

$$\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \cdots \\ w_n \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{pmatrix} \quad (8)$$

When the equation 8 is solved by a general matrix solution such as sweeping-out method, the predicted coefficients w_i are obtained. With an address of class code, the predicted coefficients w_i are stored in the memory.

Thus, with an HD signal that is true data, the learning portion can obtain the predicted coefficients w_i. The predicted coefficients w_i are stored in the memory. With any input SD signal, class information is formed. Predicted coefficients corresponding to class information are read from the memory. By linear combination of the values of SD pixels in the vicinity of the considered pixel and the predicted coefficients, the value of the considered pixel can be formed. Thus, an output HD image corresponding to any input SD image can be generated.

When the learning portion 54 determines a representative value for each class rather than a predicted coefficient, a process corresponding to a flow chart shown in Fig. 13 is performed. Step 71 "start", step 72 "forming learnt data", step 73 "data end", and step 74 "determining class" in Fig. 10 are similar to those at steps 61, 62, 63, and 64 in Fig. 12.

At step 75 "normalizing", the values of pixels are normalized. In other words, when a value (an input value) of an HD pixel is y, the input data is normalized corresponding to (y - base) / DR. In the pixel arrangement shown in Fig. 11, when one block is composed of pixels a to i, DR represents the difference (dynamic range DR) between the maximum value and minimum value of pixels in the block. base is a reference value of the block. For example, base is the minimum value of pixels of a block. base may be the average value of pixels of a block other than the minimum value. By the normalization, the relative levels of pixels can be considered.

At step 76 "determining representative value", as with the case shown in Fig. 9, the cumulative number of the class, n(c), is obtained. In addition, the representative value g(c) is obtained. In other words, the representative value g(c)' newly formed is represented as follows.

$$g(c)' = \{(y - base) / DR + n(c) \times g(c)\} / n(c + 1) \quad (9)$$

The representative value for each class obtained in such a manner is stored in the memory.

Instead of the ADRC circuit, any information compressing means for class categorization such as DCT (Discrete Cosine Transform), VQ (Vector Quantifying), or DPCM (prediction encoding) circuit can be provided.

The class categorizing adaptive processing circuit 9 learns the relation between an SD signal and an HD signal in a time region corresponding to the characteristics of the true image. Thus, the HD signal corresponding to the SD signal is generated. Since the class is adaptively selected corresponding to the level distribution of the SD signal. Thus, the up-conversion corresponding to local characteristics of the image can be performed. Unlike with the interpolating filter, an HD signal with a compensated resolution can be obtained.

Returning to Fig. 5, an impulse component 6b in the frequency region is supplied from the categorizing circuit 5. The gain converting circuit 10 as a first processing circuit compensates the resolution in the frequency region. In other words, the gain conversion compensates the decrease of the high band gain due to the signal process as shown in Fig. 14. As with the class categorizing adaptive processing circuit 9, the gain converting circuit 10 has a memory that stores a mapping table that compensates a high band. As with the above-described class categorizing adaptive processing circuit 9 in the time region, there are two types of mapping tables, one of which outputs gain converting ratio and the other of which outputs a predicted value of the gain.

Fig. 15 shows the construction of forming a mapping table of the gain converting portion 10 in the learning state. HD video data to be learnt is supplied to an input terminal 81. Thereafter, the HD video data is supplied to a sub line/sub sample circuit 82. The circuit 82 performs vertical decimation operation for sub lines and horizontal decimation operation for sub samples. Thus, the sub line/sub sample circuit 82 generates a video signal with the similar resolution to that of the SD video signal.

The sub line/sub sample circuit 82 is connected to a delay circuit 83 and a D/A converter 90. The delay circuit 83 delays the input data until the input data is categorized as a class so as to match the timing of the operation. The delay circuit 83 is connected to a block segmenting circuit 84. For example, data of a block structure composed of (4 x 4) are formed. The output of the block segmenting circuit 84 is supplied to a DCT circuit 85. The DCT circuit 85 performs cosine transformation for the input signal. The DCT circuit 85 generates coefficient data in zigzag sequence from DC component to lower order AC components to higher order AC components.

The coefficient data that is output from the DCT circuit 85 is supplied to a dividing circuit 86. The dividing circuit 86 obtains a gain converting ratio for coefficient data necessary for compensating the high band. The dividing circuit 86 outputs a gain converting ratio signal to a memory 87. The memory 87 has a plurality of memory portions that store the gain converting ratios corresponding to a plurality of DCT coefficients.

To determine the deterioration of the high band of the SD video signal due to the signal process, the SD video signal that has been converted into an analog signal by a D/A converter 90 is supplied to an analog transmission portion 91. The analog transmission portion 91 performs for example recording and reproducing processes for an analog VTR. The video signal that is output from the analog transmission portion 91 is supplied to an A/D converter 92. The A/D converter 92 converts the analog signal into a digital signal. The digital signal

is supplied to a block segmenting circuit 93.

The block segmenting circuit 93 forms digital video data with the same block structure as the output data of the block segmenting circuit 84. The output data of the block segmenting circuit 93 is supplied to a DCT circuit 94 and a class categorizing circuit 95. The coefficient data that is output from the DCT circuit 94 is supplied to a dividing circuit 86. For coefficient data with the same order, the dividing process is performed. A gain converting ratio signal with respect to coefficient data is generated by the dividing circuit 86. In other words, when a signal passes through the analog transmission portion 91, the high band frequency component is lost. The gain converting ratio signal represents how the gain (value) of each component of the DCT coefficient data varies.

Now, it is assumed that the DCT circuit 85 generates coefficient data DC and AC1 to AC15 and that the DCT circuit 94 generates coefficient data DC' and AC1' to AC5'. The dividing circuit 86 forms gain converting ratio signals G_0, G_1, ..., G_{15} by the following arithmetic operations.

$$G_0 = DC / DC', G_1 = AC / AC', \ldots, G_{15} = AC_{15} / AC_{15}'$$

In Fig. 15, by averaging a plurality of gain converting ratio signals for each coefficient (they are omitted for simplicity), the final gain converting ratio signal is obtained and stored in the memory 87.

When the gain converting ratio signal is multiplied by coefficient data of video data whose high band has been attenuated, coefficient data of the video data whose high band has been compensated can be generated. The gain converting circuit 10 shown in Fig. 5 has a memory that stores a gain converting ratio signal that has been learnt. By multiplying the coefficient data by the gain converting ratio signal, the value of the coefficient data is changed. Thus, the high band can be compensated.

The class categorizing circuit 95 categorizes the block data supplied from the block segmenting circuit 93 as a class corresponding to the level distribution thereof. For the class categorization, as described above, data compression such as ADRC is preferably performed. The class information obtained by the class categorizing circuit 95 is supplied to the memory 87 as an address in the memory. The memory 87 has a plurality of memory portions corresponding to coefficient data of DC component and AC components of all orders. The memory portions store gain conversion ratio signals corresponding to coefficient data.

An address for switching the memory portions corresponding to coefficient data is formed by an address counter 88. The address counter 88 counts the clock signal received from an input terminal 89 and generates addresses that vary one after the other. In this case, the addresses vary in synchronization with the coefficient data received from the block segmenting circuit 84. A plurality of types of HD video signals are supplied to the input terminal 81. An optimum gain converting ratio signal for each class is formed and stored in the memory 87.

Instead of the gain converting ratio, predicted values of DCT coefficients can be learnt.

The same gain converting ratio signal as that stored in the memory 87 is stored in the memory of the gain converting circuit 10 shown in Fig. 5. The output signal of the block segmenting circuit 2 is supplied to the gain converting circuit 10 that categorizes the signal as a class. The gain converting circuit 10 multiplies each component of the DCT coefficient data by the gain converting ratio signal so as to adjust the gain. Thus, the high band of the frequency region is compensated. The impulse component 6b in the frequency region is supplied to the gain converting circuit 10. When a signal composed of various components including a flat component is converted, a non-linear component is mixed and thereby the accuracy is deteriorated. Thus, the gain cannot be correctly converted. With the same reason, when the learning operation shown in Fig. 15 is performed, an impulse signal is used.

The above-described embodiment is applied for the compensation of resolution of a video signal, more particularly, for up-conversion from an SD signal to an HD signal. However, the present invention can be also applied for other than the signal process. For example, the present invention can be applied for a noise removing circuit.

As noises, impulse noise (white noise) and pink noise are generally known. The impulse noise can be removed by a median filter in a time region. In the median filter, values of three successive sample data on the time axis are compared. When the value of the middle sample data is not present between the values of the earlier and later sample data, the middle sample data is treated as a noise. The value of the middle sample data is substituted with a nearer value of the earlier and later sample data (for example, refer to Japanese Utility Model Registration Examined Publication No. HEI 3-19094).

An impulse noise in the time region can be removed by the median filter. However, a flat noise (pink noise) cannot be easily removed in the time region. In embodiments of the present invention, in addition to the noise removing process in the time region using the median filter, for example, a digital low-pass filter is used to remove the noise. Thus, the noise component that is flat in the time region can be removed. Consequently, various noises in the input digital signal can be removed.

In addition to the compensation of resolution and the noise removing circuit, the present invention can be

applied for other digital signal processes.

As described above, in embodiments of the present invention, a digital information signal is separated into a steady component and a transient component. These components are individually processed in a time region and a frequency region so that they can be properly represented. After these processes, the components are mixed. Thus, the process time can be reduced, the hardware scale can be reduced, and the accuracy of the signal process can be improved in comparison with the case that the processes in the time region and the frequency region are performed in two stages.

Having described specific preferred embodiments of the present invention with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments, and that various changes and modifications may be effected therein by one skilled in the art without departing from the scope of the invention as defined in the appended claims.

Claims

1. A digital signal processing apparatus, comprising:
   at least one analyzing means of first analyzing means for analyzing an input digital signal in a time region and second analyzing means for analyzing the input digital signal in a frequency region;
   categorizing means for categorizing the input digital signal into two components corresponding to an analyzed output of said analyzing means;
   first and second processing means for adaptively processing the signals categorized by said categorizing means in the time region and in the frequency region, respectively; and
   mixing means for mixing outputs of said first and second processing means.

2. A digital signal processing apparatus, comprising:
   analyzing means for analyzing the frequency of an input digital signal;
   separating means for separating an impulse component and a flat component in a frequency region corresponding to an analyzed output of said analyzing means;
   first processing means for receiving the impulse component of the output of said separating means and processing the impulse component in the frequency region;
   first converting means for converting the flat component of the output of said separating means into a signal in a time region;
   second processing means for processing an output of said first converting means in the time region;
   second converting means for converting an output of said first processing means into a signal in a time region; and
   mixing means for mixing an output of said second converting means and an output of said second processing means.

3. The digital signal processing apparatus as set forth in claim 1 or 2, wherein the input digital signal is a digital video signal.

4. The digital signal processing apparatus as set forth in claim 3, said first and second processing means produce a second digital video signal having resolution higher than resolution of said digital video signal.

5. The digital signal processing apparatus as set forth in claim 1 or 2, wherein said analyzing means in the frequency region includes a orthogonal transformation means.

6. The digital signal processing apparatus as set forth in claim 5, wherein the orthogonal transformation means performs either discrete cosine transformation or fast fourier transformation.

7. The digital signal processing apparatus as set forth in claim 1 or 2, wherein said processing means includes a digital filter with a limited number of taps.

8. The digital signal processing apparatus as set forth in claim 1 or 2, wherein said processing means is a noise removing circuit.

9. The digital signal processing apparatus as set forth in claim 8, wherein the noise removing circuit is constructed of a median filter or a low-pass filter.

10. A digital signal processing method, comprising the steps of:
analyzing an input digital signal in a time region or in a frequency region as an analyzing step;
categorizing the input digital signal corresponding to the analyzed result as a categorizing step;
adaptively processing the categorized signals in the time region and in the frequency region as first and second processing steps, respectively; and
mixing the results of said first and second processing steps.

11. A digital signal processing method, comprising the steps of:
analyzing the frequency of an input digital signal as a frequency analyzing step;
separating an impulse component and a flat component in the frequency region corresponding to the analyzed result as a separating step;
receiving the impulse component and processing the impulse component in the frequency region as a first processing step;
converting the flat component into a signal in a time region as a first converting step;
processing the resultant signal converted by said first converting step in the time region as a second processing step;
converting the result of said first processing step in the time region as a second converting step; and
mixing the resultant signal of said second converting step and the resultant signal of said second processing step.

12. An image converting apparatus with a compensable resolution, comprising:
at least one analyzing means of first analyzing means for analyzing a first digital image signal having a first resolution in a time region and second analyzing means for analyzing said first digital image signal in a frequency region;
categorizing means for categorizing the first digital image signal into at least two kinds of components corresponding to an analyzed output of said analyzing means;
processing means having at least first and second processing each component for adaptively processing each component categorized by said categorizing means so as to form a second digital image signal with a higher resolution than a first resolution of the first digital image signal; and
mixing means for mixing each output of said processing means.

13. An image converting apparatus with a compensable resolution, comprising:
analyzing means for analyzing the frequency of a first digital image signal with a first resolution;
separating means for separating said first digital image signal into an impulse component and a flat component in a frequency region according to the analyzed output of said analyzing means;
first processing means for receiving the impulse component from said separating means and processing the impulse component in the frequency region so as to form a second digital image signal with a second resolution higher than the first resolution;
first converting means for converting an output of said first processing means into a signal in a time region;
second converting means for receiving the flat component from said separating means and converting the flat component into a signal in the time region;
second processing means for receiving the signal in the time region from said second converting means and processing the signal in the time region so as to form a second digital image signal with a second resolution; and
mixing means for mixing an output of said first converting means and an output of said second processing means.

14. The image converting apparatus as set forth in claim 12 or 13,
wherein said analyzing means in the frequency region includes performing orthogonal transformation means.

15. The image converting apparatus as set forth in claim 14,
wherein the orthogonal transformation performs either fast fourier transformation or discrete cosine transformation.

16. The image converting apparatus as set forth in claim 12 or 13,

wherein said first processing means comprises:

class categorizing means for receiving the impulse component and determining a class corresponding to the impulse component;

compensation value generating means for generating a compensation value of the impulse component so as to obtain the second resolution that is higher than the first resolution for each determined class; and

compensating means for compensating the impulse component corresponding to the generated compensation value.

17. The image converting apparatus as set forth in claim 16,

wherein said compensation value generating means has a memory for storing the compensation value for each class, and

wherein the compensation value for each class is pre-learnt by a digital image signal with the second resolution and a digital image signal with the first resolution that is lower than the second resolution, the digital image signal with the first resolution being obtained by processing the digital image signal with the second resolution.

18. The image converting apparatus as set forth in claim 17,

wherein the impulse component is used in the frequency region as data to be learnt.

19. The image converting apparatus as set forth in claim 17,

wherein the digital image signal with the second resolution is processed by an analog processing portion so as to attenuate the high band component of the digital image signal, obtain the ratio of a component of which the second digital image signal has been converted into the frequency region and a component of which the digital image signal with the high band component attenuated has been converted into the frequency region, and obtain the compensation value for each class corresponding to the ratio when the compensation value for each class is learnt.

20. The image converting apparatus as set forth in claim 12 or 13,

wherein said first processing means comprises:

class categorizing means for receiving the impulse component and determining a class corresponding to the impulse component; and

compensation value generating means for generating the value representing the impulse component with the second resolution that is higher than the first resolution for each determined class.

21. The image converting apparatus as set forth in claim 20,

wherein said compensation value generating means has a memory for storing a value representing the impulse component with the second resolution for each class, and

wherein the value representing the impulse component with the second resolution for each class is pre-learnt with the digital image signal with the second resolution and the digital image signal with the first resolution that is lower than the second resolution, the digital image signal with the first resolution being obtained by processing the digital image signal with the second resolution.

22. The image converting apparatus as set forth in claim 21,

wherein the impulse component is used in the frequency region as data to be learnt.

23. The image converting apparatus as set forth in claim 21,

wherein the digital image signal with the second resolution is processed by an analog processing portion so as to attenuate the high band component of the digital image signal, obtain the ratio of a component of which the second digital image signal has been converted into the frequency region and a component of which the digital image signal with the high band component attenuated has been converted into the frequency region, and obtain the compensation value for each class corresponding to the ratio when the compensation value for each class is learnt.

24. The image converting apparatus as set forth in claim 12 or 13,

wherein said second processing means comprises:

class categorizing means for categorizing a considered pixel as a class with a plurality of pixels of the signal in the time region received from said second converting means, the plurality of pixels being

present spatially and/or temporally adjacent to the considered pixel;

predicted coefficient generating means for generating a predicate coefficient for each determined class; and

estimating means for estimating a predicted value by linear combination of values of a plurality of pixels of the signal in the time region received from said second converting means and the predicted coefficient.

25. The image converting apparatus as set forth in claim 24,

wherein said predicted coefficient generating means has a memory for storing a predicted coefficient for each class, and

wherein a predicted coefficient for each class is pre-learnt so that the error between the generated value of the considered pixel and the true value thereof is minimized when the value of the considered pixel is generated by linear combination of the values of a plurality of pixels spatially and/or temporally adjacent to the considered pixel and the predicted coefficient.

26. The image converting apparatus as set forth in claim 12 or 13,

wherein said second processing means comprises:

class categorizing means for categorizing a considered pixel as a class with a plurality of pixels of the signal in the time region received from said second converting means, the plurality of pixels being present spatially and/or temporally adjacent to the considered pixel; and

predicted value generating means for generating a predicted value for each determined class so as to generate the value of the considered pixel.

27. The image converting apparatus as set forth in claim 26,

wherein said predicted value generating means has a memory for storing a predicted value for each class, and

wherein the value of which the cumulative value of the value obtained for each class is divided by the number of cumulative times is stored in the memory as the predicted value for each class.

28. The image converting apparatus as set forth in claim 27,

wherein said predicted value generating means has a memory for storing a predicted value for each class,

wherein a block composed of a plurality of pixels including the considered pixel is formed when learnt,

wherein a value of which a reference value of the block is mixed from the value of the considered pixel corresponding to the dynamic range of the block, and

wherein a value of which the cumulative value of the normalized values is separated by the number of cumulative times is stored in the memory as the predicted value for each class.

29. An image converting method with a compensable resolution, comprising the steps of:

analyzing a first digital image signal having first resolution in a time region or in a frequency region as an analyzing step;

categorizing the first digital image signal corresponding to the analyzed result;

adaptively processing the categorized signals in the time region and in the frequency region so as to form a second digital signal with a second resolution that is higher than the first resolution (as first and second processing steps); and

mixing the each result of said first and second processing steps (as a mixing step).

30. An image converting method with a compensable resolution, comprising the steps of:

analyzing the frequency of a first digital image signal with a first resolution (as a frequency analyzing step);

separating an impulse component and a flat component in a frequency region corresponding to the analyzed result (as a separating step);

receiving the impulse component and processing the impulse component in the frequency region so as to form a second digital image signal with a second resolution that is higher than the first resolution (as a first processing step);

converting a signal as the result of said first processing step into a signal in a time region (as a first converting step);

converting the flat component into a signal in the time region (as a second converting step);

processing a signal as the result of said first converting step in the time region so as to form the second digital image signal with the second resolution that is higher than the first resolution (as a second processing step); and

mixing the signal as the result of said first converting step and the signal as the result of said second processing step (as a mixing step).

Fig. 1A

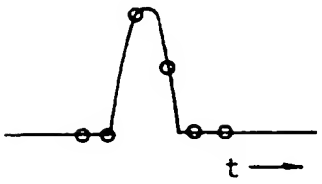


Fig. 1B

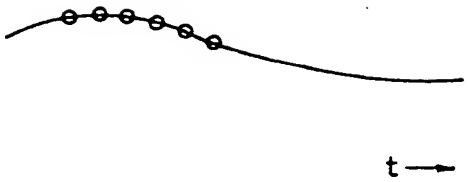


Fig. 2A

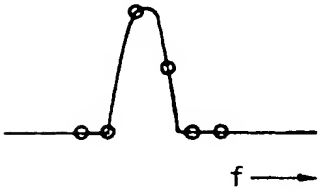


Fig. 2B

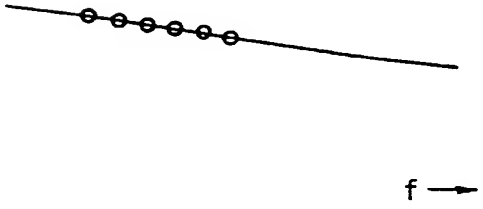


Fig. 3

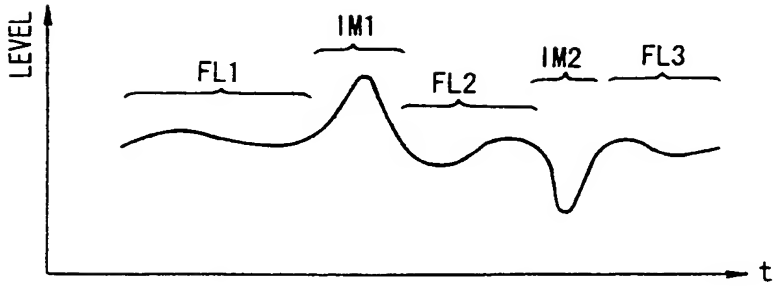


Fig. 4A

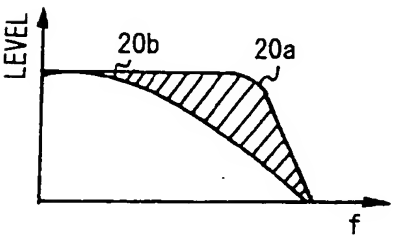


Fig. 4B

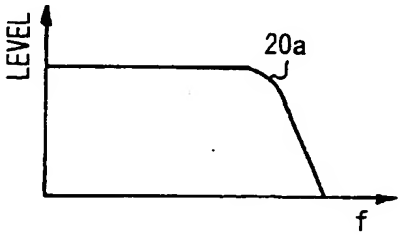


Fig. 5A

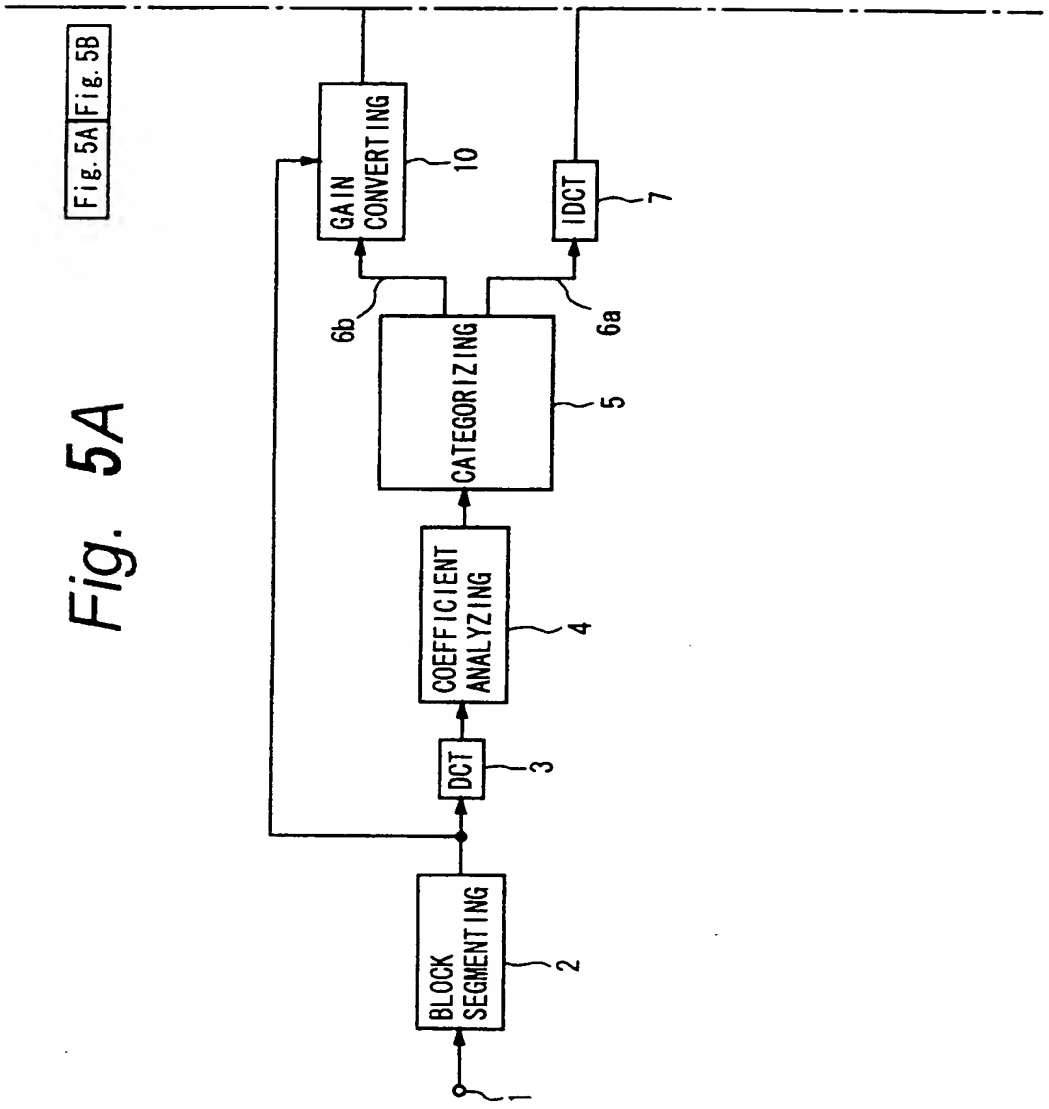


Fig. 5A
Fig. 5B
GAIN CONVERTING
10
IDCT
7
6b
6a
CATEGORIZING
5
COEFFICIENT ANALYZING
4
DCT
3
BLOCK SEGMENTING
2
1

Fig. 5B

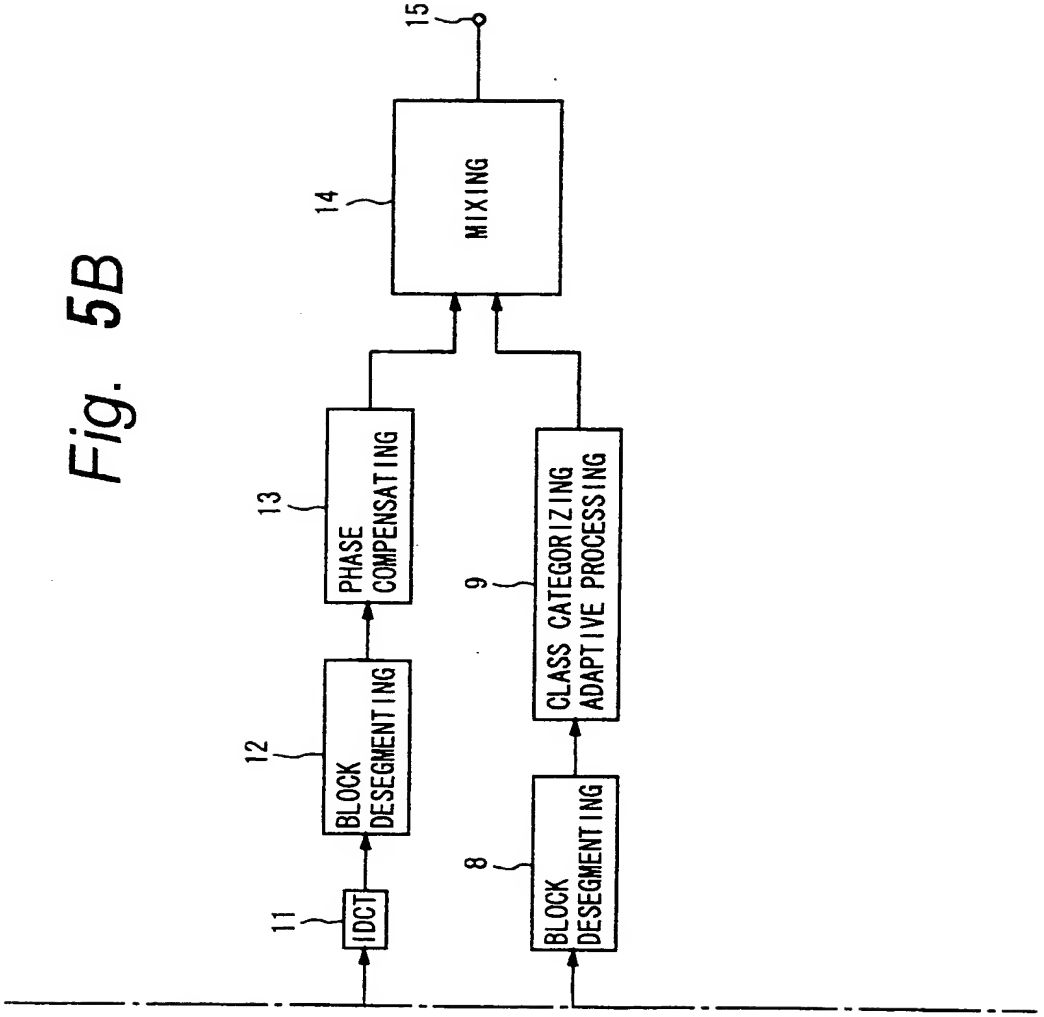

15
MIXING
14
PHASE COMPENSATING
13
BLOCK DESEGMENTING
12
IDCT
11
CLASS CATEGORIZING ADAPTIVE PROCESSING
9
BLOCK DESEGMENTING
8

Fig. 6

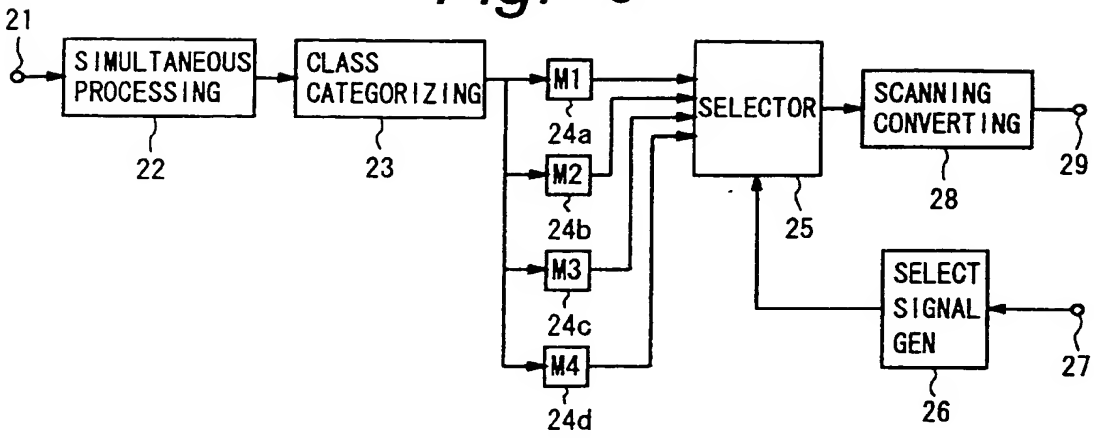

21
SIMULTANEOUS PROCESSING
22
CLASS CATEGORIZING
23
M1
24a
M2
24b
M3
24c
M4
24d
SELECTOR
25
SCANNING CONVERTING
28
29
SELECT SIGNAL GEN
26
27

Fig. 8

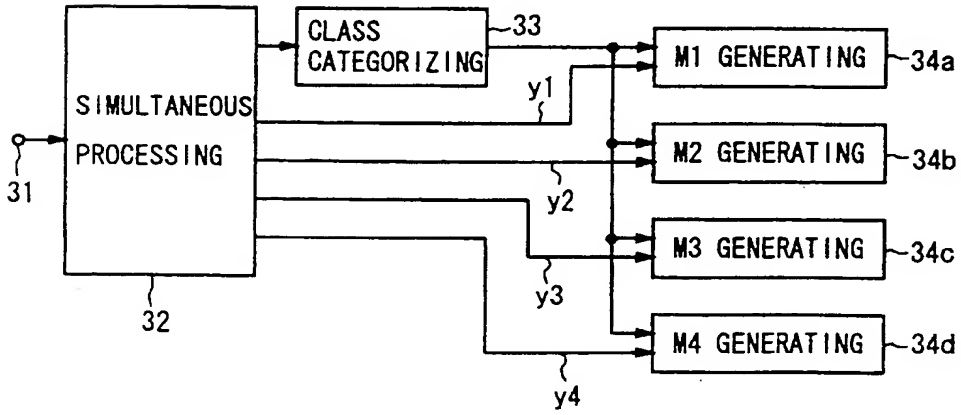

SIMULTANEOUS PROCESSING
31
32
CLASS CATEGORIZING
33
y1
y2
y3
y4
M1 GENERATING
34a
M2 GENERATING
34b
M3 GENERATING
34c
M4 GENERATING
34d

Fig. 7

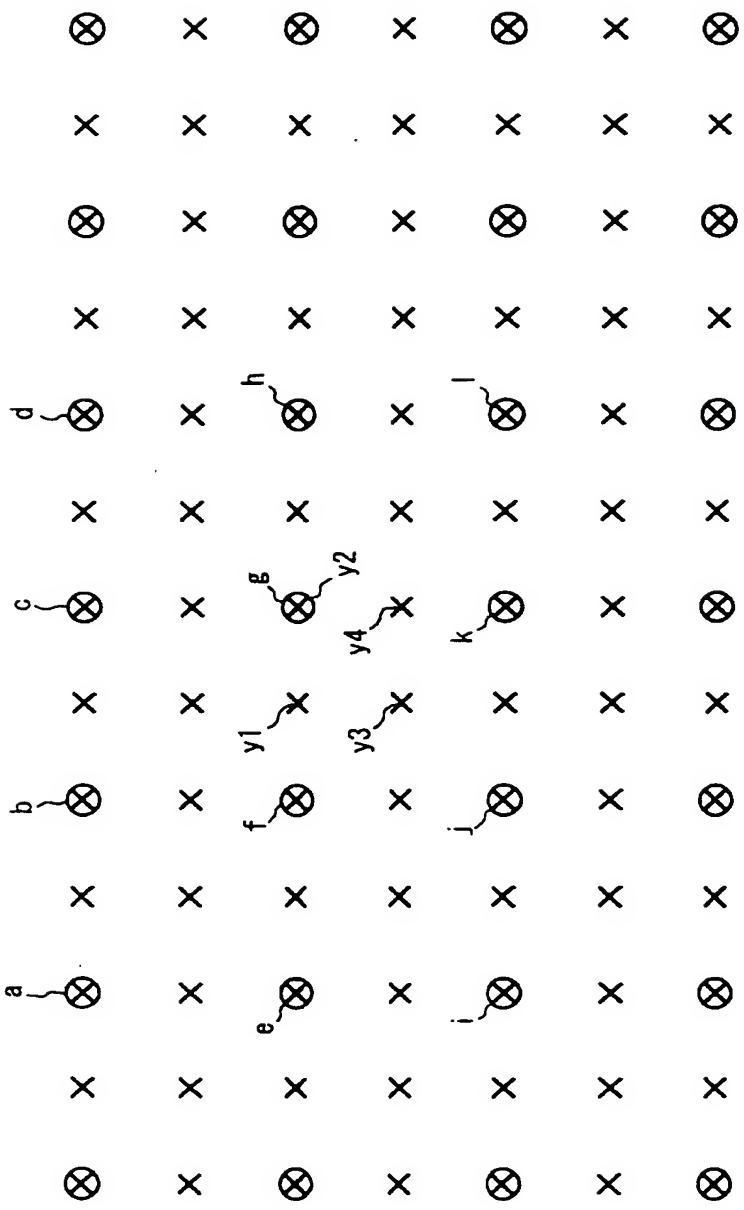


Fig. 9

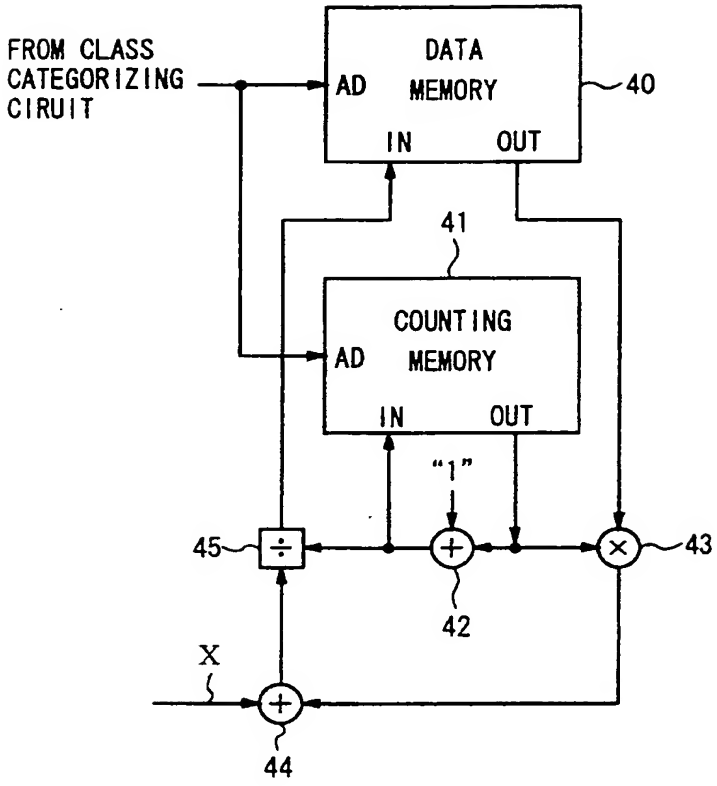

FROM CLASS CATEGORIZING CIRUIT
DATA MEMORY
AD
IN
OUT
40
41
COUNTING MEMORY
AD
IN
OUT
"1"
45
43
42
X
44

Fig. 10

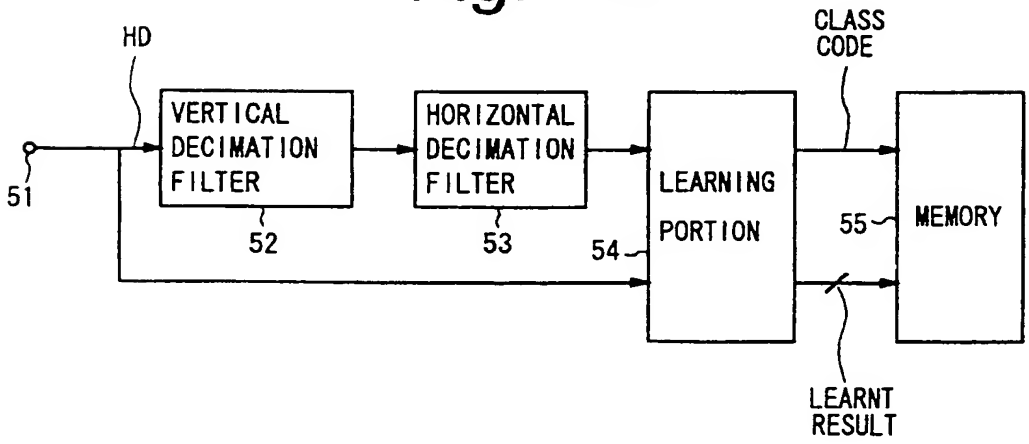

HD
CLASS CODE
VERTICAL DECIMATION FILTER
HORIZONTAL DECIMATION FILTER
LEARNING PORTION
MEMORY
51
52
53
54
55
LEARNT RESULT

Fig. 11

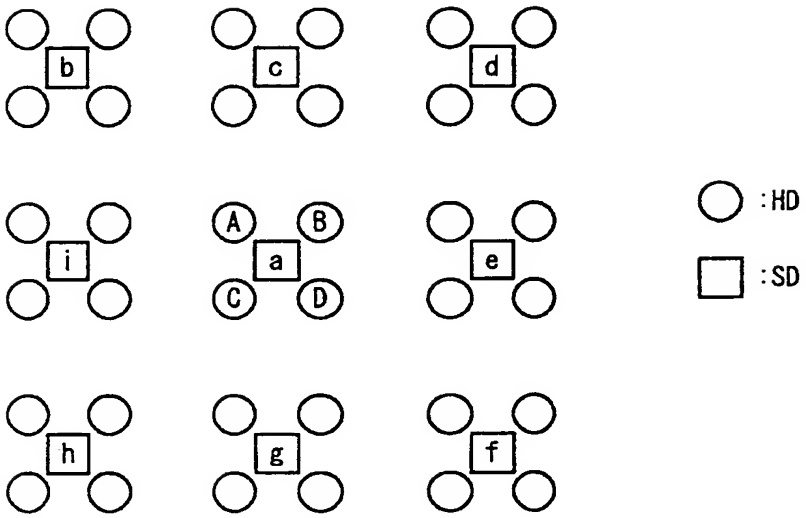


Fig. 14

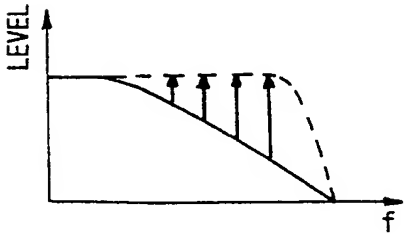


Fig. 12

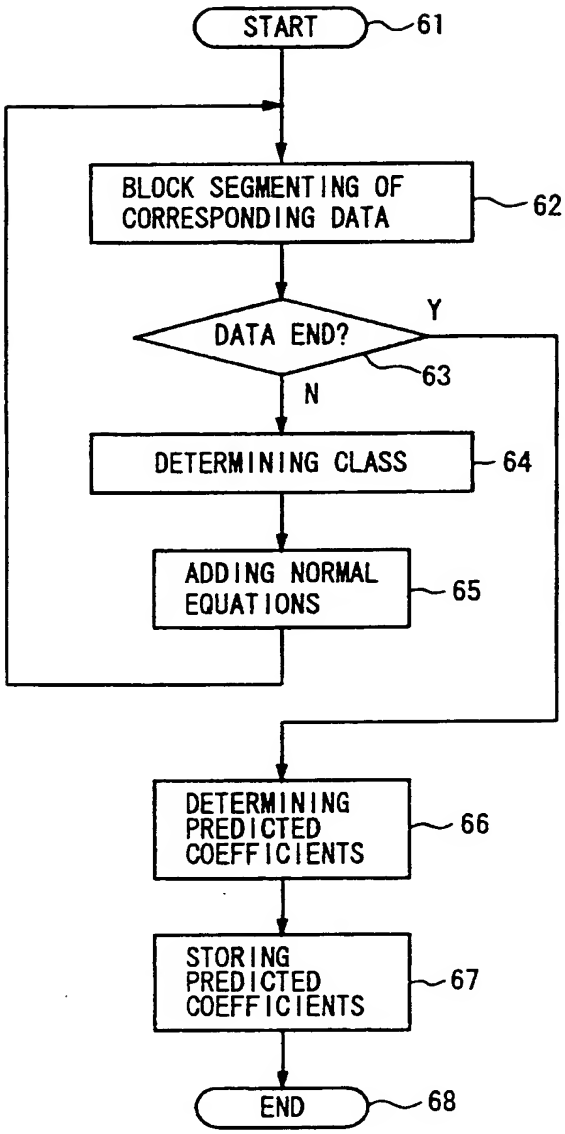

START
61
BLOCK SEGMENTING OF CORRESPONDING DATA
62
DATA END?
Y
63
N
DETERMINING CLASS
64
ADDING NORMAL EQUATIONS
65
DETERMINING PREDICTED COEFFICIENTS
66
STORING PREDICTED COEFFICIENTS
67
END
68

Fig. 13

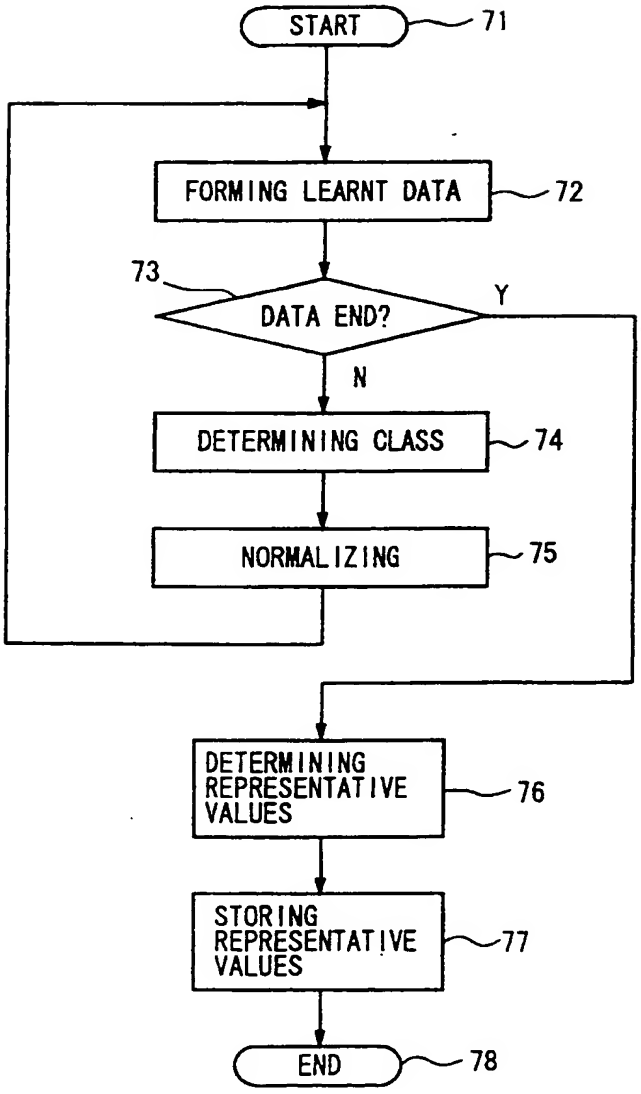
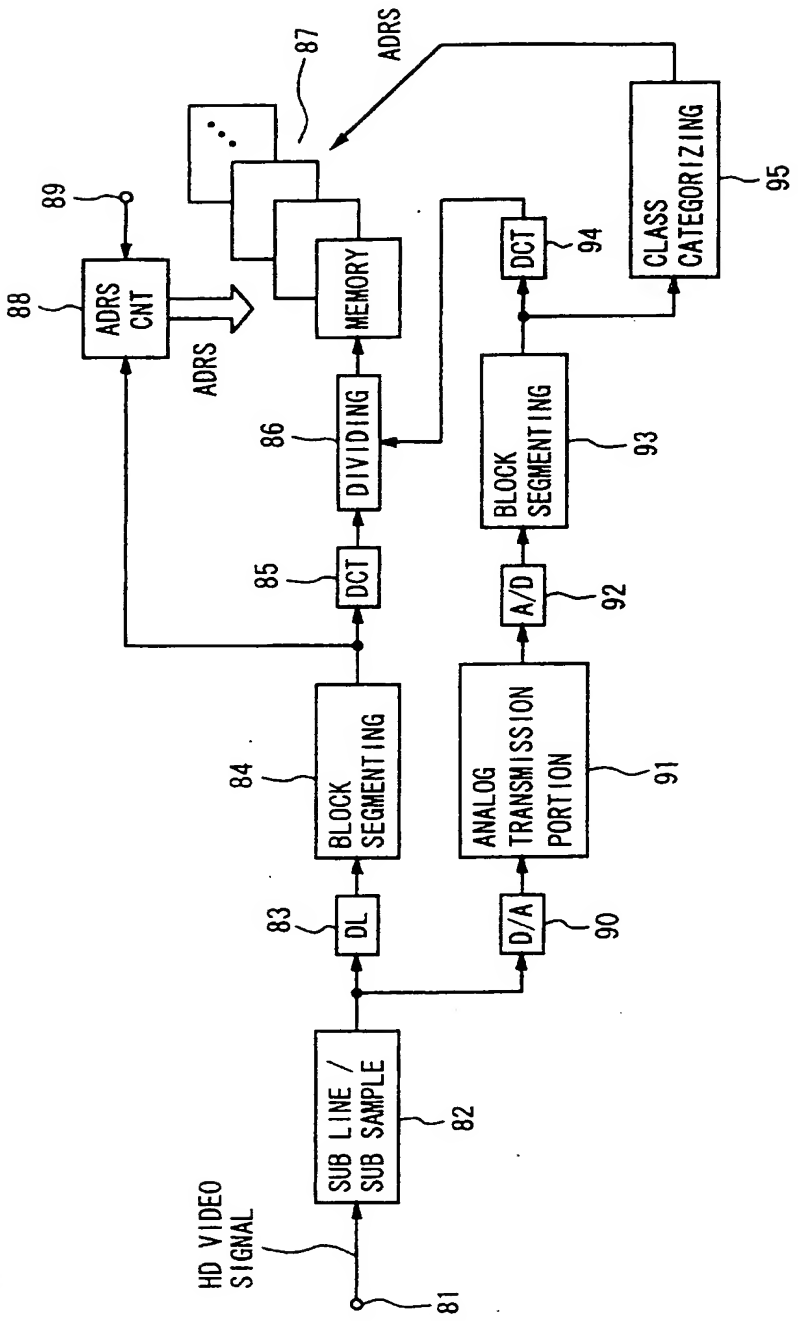

START
71
FORMING LEARNT DATA
72
73
DATA END?
Y
N
DETERMINING CLASS
74
NORMALIZING
75
DETERMINING REPRESENTATIVE VALUES
76
STORING REPRESENTATIVE VALUES
77
END
78

Fig. 15


HD VIDEO SIGNAL
81
SUB LINE / SUB SAMPLE
82
DL
83
BLOCK SEGMENTING
84
DCT
85
DIVIDING
86
87
ADRS
MEMORY
ADRS
ADRS CNT
88
89
D/A
90
ANALOG TRANSMISSION PORTION
91
A/D
92
BLOCK SEGMENTING
93
DCT
94
CLASS CATEGORIZING
95